ESD-TR-82-052

Semiannual Technical Summary
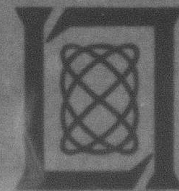
Packet Speech Systems Technology

31 March 1982

Prepared for the Defense Advanced Research Projects Agency
under Electronic Systems Division Contract F19628-80-C-0002 by
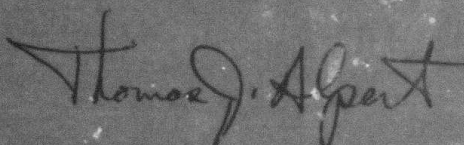
Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LEXINGTON, MASSACHUSETTS

Approved for public release; distribution unlimited.

DTIC
SELECTE
OCT 19 1982
S
D
F

82 10 18 110

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Thomas J. Alpert

Thomas J. Alpert, Major, USAF
Chief, ESD Lincoln Laboratory Project Office

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

PACKET SPEECH SYSTEMS TECHNOLOGY

SEMIANNUAL TECHNICAL SUMMARY REPORT
TO THE
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

1 OCTOBER 1981 — 31 MARCH 1982

ISSUED 18 AUGUST 1982

LEXINGTON   MASSACHUSETTS

ABSTRACT

This report describes work performed on the
Packet Speech Systems Technology Program
sponsored by the Information Processing
Techniques Office of the Defense Advanced
Research Projects Agency during the period
1 October 1981 through 31 March 1982.

Accession For

NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	

By
Distribution/
Availability Codes

Dist	Avail and/or Special
A	

TABLE OF CONTENTS

INTRODUCTION AND SUMMARY

The long-range objectives of the Packet Speech Systems Technology
Program are to develop and demonstrate techniques for efficient digital
speech communications on networks suitable for both voice and data, and to
investigate and develop techniques for integrated voice and data
communication in packetized networks, including wideband common-user
satellite links. Specific areas of concern are: the concentration of
statistically fluctuating volumes of voice traffic, the adaptation of
communication strategies to varying conditions of network links and traffic
volume, and the interconnection of wideband satellite networks to terrestrial
systems.

Previous efforts in this area have led to new vocoder structures for
improved narrowband voice performance and multiple-rate transmission, and to
demonstrations of conversational speech and conferencing on the ARPANET and
the Atlantic Packet Satellite Network.

The current program has two major thrusts: i.e., the development and
refinement of practical low-cost, robust, narrowband, and variable-rate
speech algorithms and voice terminal structures; and the establishment of an
experimental wideband satellite network to serve as a unique facility for the
realistic investigation of voice/data networking strategies. This report covers work in the following areas: compact vocoder
development based on digital LSI technology; development of Packet Voice
Terminal (PVT) and local access network (LEXNET) facilities for experiments
in packet voice; development and experimental application of a flexible
internet stream gateway (the miniconcentrator) for packet voice; planning,
coordination, and execution of multi-user packet speech experiments on the
experimental wideband satellite network (WB SATNET); and design and
development of a system for voice control of network voice conferencing in
the wideband system.

A single-card 2.4-kbps PVT-compatible Linear Predictive Coding (LPC)
vocoder based on a now-available commercial Signal-Processing Interface (SPI)
chip has been developed, constructed, and demonstrated using EPROM versions

of the SPI chips. Mask-programmed versions of the SPI chips have been ordered. The vocoder is composed of 16 integrated circuits on an 18-in.-IC area, and consumes 5.5 W.

The ten PVT units and four LEXNET/Concentrator Interface units have been used in a variety of packet speech tests in the wideband network environment. Key tests have included demonstration of two simultaneous PCM calls between PVTs on LEXNETs at Lincoln and the Information Sciences Institute (ISI), and demonstration of interoperability between a LEXNET PVT and a telephone on an emulated circuit switch. A PVT information package for the purpose of technology transfer has been distributed, and a number of industrial expressions of interest have been collected. PVT software has been expanded to include rate-control for the variable-rate embedded CVSD (ECVSD) voice coders, and additional measurement facilities for loop delay measurements in the wideband net. PVT software for internet voice conferencing has been designed and partially implemented.

The miniconcentrator now includes all basic internet protocol (IP) and point-to-point stream (ST) protocol functions in operational form, and an ST conference-handling capability has been added. The software is modularized for easy generation of versions supporting simultaneous connections to N nets, and "3-headed" gateways versions have been used routinely. Packet aggregation has been implemented to reduce the packet load on the WB SATNET, and an ST source routing mechanism has been implemented to facilitate controlled testing of a variety of network routes. Interoperability between the miniconcentrator and the Bolt, Beranek, and Newman (BBN) Voice Funnel has been demonstrated in a variety of tests. Lincoln has continued to play an active role in the coordination of activities aimed at integration and checkout of the WB SATNET. In particular, we have coordinated several cross-net speech tests involving the Lincoln, ISI, and Defense Communications Engineering Center (DCEC) WB SATNET sites.

System design, hardware procurement, and initial software development have been carried out for a voice-controlled operator (VCOP) for conference call setup in the wideband system. This voice-control facility will be interfaced to the packet voice conferencing system currently being implemented.

PACKET SPEECH SYSTEMS TECHNOLOGY

I. COMPACT VOCODER DEVELOPMENT

The programs for the three Nippon Electric Company (N.E.C.) µPD7720 Signal-Processing Interface (SPI) microcomputers in the single-board PVT-compatible LPC vocoder have been written and debugged. An order has been placed with N.E.C. Electronics, U.S.A. for factory-programmed ROM µPD7720s. Three distinct ROM contents have been specified corresponding to the LPC analyzer, pitch detector, and synthesizer. One hundred and fifty units of each of the three types have been ordered for a total of 450 parts. The necessary ROM code bit maps have been delivered to N.E.C. and delivery of approximately ten pre-production SPI LPC vocoder sets for Lincoln's inspection is expected in late May.

Initial debugging of the programs for the three SPI devices was carried out by downloading these programs into the EVAKIT-7720, the SPI hardware emulator. The performance of the programs was compared directly to the real-time LPC implementation on the LDSP, a technique which proved invaluable in debugging the SPI microcode. Six units of the EPROM version of the SPI, designated as the "µPD7720," were then obtained. The µPD7720 internal program and coefficient memories are erased under exposure to UV light and are programmed using the EVAKIT-7720. Before delivering the vocoder SPI bit maps to N.E.C., a final verification of the LPC vocoder microcode was obtained by the simultaneous operation of the three µPD7720s programmed as the LPC analyzer, pitch detector, and synthesizer with the LDSP executing the INTEL 8085 control computer tasks.

The SPI pitch detector, originally designed to function only at sampling frequencies of 8 and 10 kHz, has been made more flexible by allowing operating at arbitrary sampling frequency within the real-time constraint. These modifications have been incorporated into the factory-programmed ROM order.

Since the LPC vocoder SPI programs are intended for a factory-programmed ROM realization, a very flexible design was desired. This has been achieved

by including in all three SPI programs an initialization sequence during which key operation parameters are downloaded from the control microprocessor. These flexibility options have been chosen so that the resulting LPC SPI chip set will operate using a variety of formats (including the DoD standard format) for 2.4-kbps LPC speech communication as well as in a variety of speech recognition front-end and speech synthesis applications. The options allow choice of model order (0 to 15), choice of analog or digital pre- and de-emphasis, choice of frame size, as well as choice of linear or μ-255 law speech input and output. The pitch detector chip may be run at sampling rates of either 8 or 10 kHz. A silence threshold is provided to the pitch detector at initialization to optimize its performance for a given level of background noise on the speech input. Although it was originally expected that both LPC analyzer and synthesizer programs would fit in a single SPI program ROM, the inclusion of the initialization options has lengthened the code so that separate analyzer and synthesizer ROM masks are needed. Therefore, a total of three distinct LPC chips will be specified.

The memory usage of the three SPI programs for the compact LPC vocoder is given in Table I. The foreground and background computation times for the three chips are given in Tables II and III, respectively.

TABLE I			
LPC VOCODER MEMORY REQUIREMENTS (10th Order Model)			
	Percentage of		
	Data RAM (128 x 16)	Program ROM (512 x 23)	Data ROM (512 x 13)
Gold Pitch Detector	94	95	10
Linear Predictive Analysis	70	80	3
Synthesis	63	67	0

2

TABLE II		
LPC VOCODER FOREGROUND PROCESSING (10th Order Model)		
	Time (µs)	
	Nominal	Worst Case
Gold Pitch Detector	20	88
Linear Predictive Analysis	63	75
Synthesis	62	82

TABLE III	
LPC VOCODER BACKGROUND PROCESSING (10th Order Model)	
	Time (ms)
Gold Pitch Detector	2.0
Linear Predictive Analysis	1.7
Synthesis	0.2

The software for the single-board vocoder's INTEL 8085 control microcomputer has also been written and debugged. The control microcomputer tasks are: (a) frame synchronization for the three SPI processors; (b) 2400-bps coding of the analysis parameters from the '7720 analyzer and pitch detector, and transmission of the coded parameters to the protocol processor; and (c) decoding of synthesis parameters from the protocol processor and transmission to the synthesizer SPI. The control microcomputer also initializes the three SPIs with the vocoder parameters desired. The PVT compatible board is currently using a 130-µs sampling interval and a 158 sample frame size to be interoperable with the ARPA LPC impler ˈtations. The control microcomputer code also contains an adaptive threshol ᵢilence detector providing silence information to the protocol processor. ᵥₑ INTEL 8085 control microcomputer requires 2013 bytes of PROM and bytes of RAM. With an 8-MHz clock, the INTEL 8085 program requires 15 ms ⸱ 20.5-ms speech frame or equivalently 73 percent of real-time.

A second vocoder board has been wire-wrapped resulting in a total of two LPC vocoders. Using the six µPD7720s, three for each vocoder, and the INTEL 8085 control microcomputers, the two single-board LPC vocoders have been installed in the PVT vocoder slot and successfully demonstrated in communicating over the LEXNET. While some fine-tuning remains to be performed on the audio subsystem, the vocoder pair has been found quite intelligible in actual conversations over the LEXNET.

Manufacturing of the AMI S3505 CODEC-with-filters chip, which is used in the single-board vocoder, has been discontinued by AMI. AMI has replaced the device with the functionally equivalent AMI S35057A. There are several advantages of the newer S35057A: the S35057A consumes less power than the S3505 and includes two undedicated operational amplifiers thereby eliminating the need for an op-amp package on the vocoder board. The S35057A also inputs and outputs the analog speech signals at a higher voltage than the S3505 which may result in a better signal-to-noise ratio on the device itself. However, the S35057A dₑ s differ from the S3505 in pin designations and some electrical specifications. The single-board vocoder therefore has been

4

redesigned to accommodate the S35057A. A block diagram of this design is shown in Fig. 1. A photograph depicting the major functional units of the vocoder is shown in Fig. 2.

Finally, while the original estimate of power consumption for the single-board LPC vocoder was 8.6 W, in actual use the measured power consumption was found to be only 5.5 W.

Lincoln is considering the development of a stand-alone version of the PVT-compatible LPC vocoder later in the fiscal year. This board would provide the basic hardware capability to support a variety of applications within the NSC community including recognition and voice response. The vocoder would be equipped with self-contained audio, parallel and RS-232 type interfaces, and an 8085A-2 control processor capable of supporting a generous memory complement. A document has been prepared in collaboration with ISI outlining the capabilities of the naked N.E.C. chips, the conceptual design of the stand-alone LPC engine, and a suggested protocol for operating the RS-232 serial link. This document has been circulated to the ARPA Packet Speech community for comment and critique.
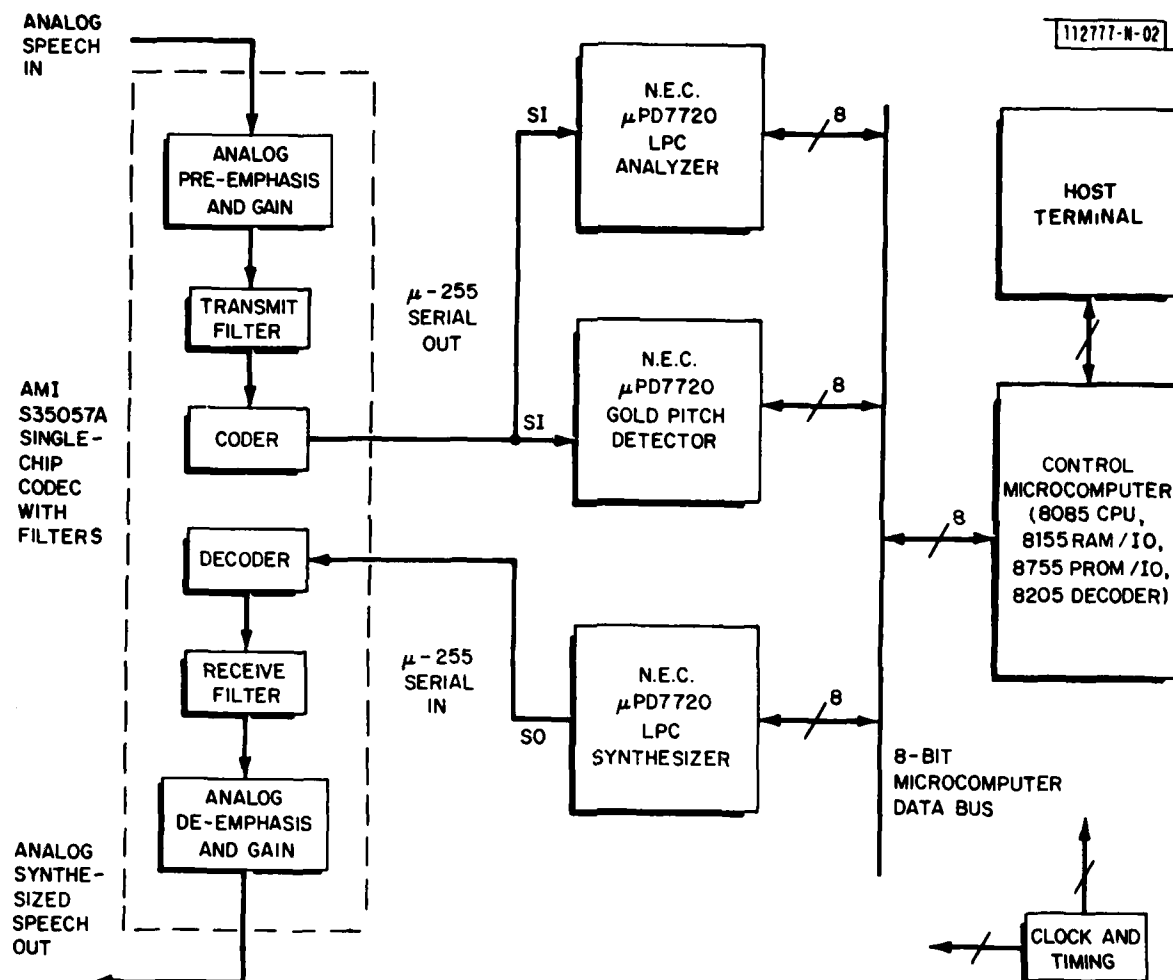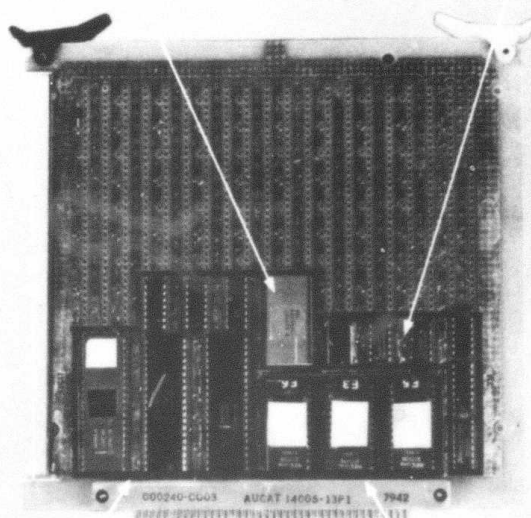
Fig. 1. Lincoln compact LPC vocoder block diagram.

Fig. 2. Compact LPC vocoder.

II. PACKET VOICE TERMINAL

A. PVT and LEXNET HARDWARE STATUS

The PVT and LEXNET equipment continues to receive heavy use in a variety of packet speech tests in the wideband network environment. There are currently ten PVTs and four LEXNET/Concentrator Interfaces (LCIs) in operation. Currently there are six PVTs and two LCIs at Lincoln. Five of the PVTs and the LCIs are dedicated to the support of the channel experiments and to the development of conferencing software. The sixth terminal has been used for the development and testing of the LPC vocoder, the ECVSD vocoder, and a timer card which will be used for making network timing measurements. Two PVTs and an LCI remain at ISI. A key milestone, achieved for the first time in mid-November 1981, was the demonstration of two simultaneous 64-kbps PCM calls over the WB SATNET between PVTs on LEXNETs at Lincoln and ISI. One PVT and one LCI have been shipped to SRI to help in testing their gateway and satellite channel equipment. BBN currently has one PVT and one LCI for Voice Funnel testing. Two new protocol processor PROM cards have been built, bringing to four the number of PVTs which can be used in stand-alone configuration. For both SRI and BBN, a new protocol code has been entered into the PROMs on the protocol processors, allowing a terminal to transmit both protocol and speech packets to itself, looping locally through the gateway or at the satellite. The single PVT and LCI combination thus allows both SRI and BBN to fully exercise their systems. After a period of Voice Funnel testing, the units currently at BBN will be shipped temporarily to DCEC to allow demonstration of four-site packet speech at the June packet speech meeting.

After the initial November demonstration of packet speech over the channel to ISI, a number of additional milestone experiments were carried out. Of particular note was a speech experiment from a PVT at Lincoln through a LEXNET and Voice Funnel, over the channel to PVT at ISI (through a miniconcentrator and LEXNET), achieved for the first time on 3 February. In addition, packet/circuit interoperability was demonstrated on 28 January by means of a call over the satellite between a LEXNET PVT and an emulated circuit switch connected to the PSAT via a packet/circuit gateway.

9

The fourth LCI unit, which has been constructed and tested during this reporting period, has a different configuration from the previous units, providing separate hardware for packets going to and from the concentrator. The new LCI unit consists of 5 cards: one modem card, two buffer control cards, and two LCI cards. Completely independent hardware and buffering is thus provided for the two directions of packet transmission. This configuration should provide higher throughput by eliminating the hardware sharing, and throughput tests are currently being conducted in conjunction with the UMC-Z80 in the miniconcentrator.

A new version of the buffer control processor PROM program has been installed in the PVTs. This new version changes the way conference packets are handled in order to reduce the processing burden on the protocol processor. Originally, all conference packets were addressed to a "broadcast" address which was accepted by all PVTs. Thus, a terminal would receive packets from all participants in all conferences including itself. The new version of the buffer control processor code filters these packets to reduce the burden on the protocol processor. The identity of the conference to which a broadcast packet belongs is indicated by an extension byte in the local net header. The buffer control processor matches this field against a mask and discards packets not belonging to conferences of interest. It also checks the source address field and discards its own transmitted packets.

B. PACKET VOICE TERMINAL TECHNOLOGY TRANSFER

An information package describing the PVT and requesting "expressions of interest" in technology transfer from commercial vendors was sent to 15 potential vendors, based on a telephone survey identifying potential respondents to a future request-for-proposals (RFP).

The package described the current PVT configuration and support facilities and proposed three possible options for outside replication: (1) direct electrical and mechanical replication of the existing hardware; (2) repackaging the existing hardware (perhaps using PC cards), but retaining the basic electrical design; and (3) re-engineering the entire system to meet function and protocol specifications. This last option would allow the

10

incorporation of significant design improvements using more recent technology such as the 10-Mbps ETHERNET.

Eleven responses have been received. All but one of these, Singer-Kearfott, expressed some interest in the program. Five contractors, Codex, ECI Division of E Systems, Litton Amecom, Adams-Russell, and ITT have visited Lincoln to get more information. One contractor, Time and Space Systems, has provided some budgetary estimates for the repackaging options. Two others, GTE and Telesystems have asked to receive the RFP when issued but are offering no technical suggestions or budget estimates at this time. Texas Instruments has indicated that it is in the process of preparing a response.

DARPA has indicated that actual procurement of PVTs will be deferred until FY 83 and that a minimal cost option is desired. The detailed procurement specification, to be written later in this fiscal year, will therefore aim at essentially a direct copy of the existing units, with possible small changes aimed at easing production and checkout. Of the companies responding to the expression-of-interest, only Adams-Russell expressed strong interest in the direct-copy option. The other companies, while not ruling out the direct-copy option, were most interested in options which would allow them more flexibility in engineering the package.

C. PACKET VOICE TERMINAL SOFTWARE DEVELOPMENT

1. Embedded Coding Rate Control Protocols

PVT protocol software to support a variety of data packaging and rate control options, as described in the last Semiannual,* has been implemented and tested successfully on a LEXNET with a pair of embedded CVSD-based waveform coders. The new PVT software can dynamically change voice transmission rate and the amount of data in each packet. Currently, the talker uses the touchtone pad to key in the desired options, which can be changed in mid-talkspurt. The two PVTs _n communication need not transmit with the same rates and packet sizes. Packets are prioritized so that a gateway can discard lower priority packets when it is unable to handle a

---

* Packet Speech Systems Technology Semiannual Technical Summary, Lincoln Laboratory, M.I.T. (30 September 1981), DTIC AD-A112045.

11

higher transmission rate. Later, protocols will be developed so that
gateways can send messages to PVTs requesting a drop in transmission rate and
so that a PVT can make a request to the gateway to allow transmission at a
higher rate.

2.   Conferencing

The packet voice conferencing capability now being implemented involves
extensions to both the ST and NVP protocols. ST is being extended to provide
multi-address distribution of conference data and control packets. NVP is
being extended to carry out conference setup and takedown as well as the
sending of conference data packets.

There are two control mechanisms involved in a voice conference. One is
called "access control" and is concerned with controlling which speech hosts
(PVTs) are to be allowed to participate in any particular conference. The
other is called "floor control" and is the mechanism that decides which
participant is able to speak and to the conference at any point in time. In
our design, access control is provided by a program called the Access
Controller (AC) that resides in a host in the internet. Floor control is
carried out by a distributed control algorithm that resides in each
participant's PVT.

At the protocol level, conferences are carried out in a "meet me" style.
Each participant contacts the AC to get permission to enter the conference.
IF the participant is on a list of acceptable hosts, the AC returns an ACCEPT
message with detailed information about the conference. The list of
acceptable participants is assumed to have been provided in advance by some
person acting as a "conference originator." The AC allows open participation
(anyone can join) as an option. In our first implementation, only open
participation is offered.

When a participant enters a conference, the AC gives a participant
number to the entering PVT, and the PVT begins a process of establishing an
ST conference connecti to all other participating PVTs with lower
participant numbers. The AC provides the network address of these
participants upon request by either PVTs or gateways that become involved in
the conference connection.

12

When two or more participants have established contact, speech can flow on the conference connection. Our floor control algorithm allows full-duplex communication when only two participants are present, but changes to half-duplex when a third enters, since the PVT can produce only one speech signal for its user. In the half-duplex mode, a PVT will start sending speech only when there has been a short period of silence (no speech received from the network), and the user goes above the speech activity threshold. Since more than one participant may attempt to talk at the same time, contention for the floor can result and is taken care of by assigning a precedence to each participant (the same as the order of conference entry in our first implementation). The PVT will stop transmitting if speech is received from a higher precedence participant or if a special "I AM TALKING" control packet is received from such a participant. The control packet is given a different type of service by the ST agents in the network to increase its probability of making it through the network relative to that of the speech packets which may be constrained to a resource allocation that is insufficient to pass all the data packets that could arise in such a floor contention situation.

An initial simplified version of conferencing, including an Access Controller, is written and seems to be working correctly. A PVT is serving as the internet host for the initial implementation of the AC. Conferences have been held among three PVTs on two LEXNETs. Conferences between two PVTs on the same net have been set up and taken down successfully. Many more debugging sessions are needed to exercise all aspects of the protocol and to ascertain that this code is totally reliable.

Several features remain to be added. The floor conrol algorithm has been designed, but not yet implemented. A time-out and retransmission mechanism is needed to ensure reliable communication of control messages which have been properly acknowledged in a given period of time. Finally, the PVT conferencing code must be adapted to handle a variety of vocoders, including the new 2400-bps LPC vocoder on the N.E.C. chips.

3.   PVT Measurement Facilities

Additional features have been added to the PVT software to allow a wider range of experiments to be performed at a single site. PVT extension ONE has

13

been designated as the ECHO extension. When a PVT receives a call set up
message directed to the ECHO extension it completes protocol but does not
ring the phone or turn on the vocoder. Incoming speech is echoed back to the
sender. This allows one site to conduct cross-country tests or demos without
disturbing another site. This feature was utilized by ISI when it was after
closing time on the East Coast. A limited ability to check for channel
errors has also been added. When this feature is used the PVT sends a series
of packets containing known bit patterns. The PVT also checks incoming
packets against the known bit patterns and records the number of bytes
received in error. Again the callee's phone is not rung and the vocoder is
not activated. This error test feature can be used in conjunction with the
ECHO option to conduct loop tests over the channel from a PVT.

D. PVT-BASED NETWORK MEASUREMENTS

A series of delay histogram measurements were taken using a loop
configuration as shown in Fig. 3. Packets were sent from a PVT on a LEXNET
to itself, looping back from various points in the WB network. The
experiments made use of the standard PVT point-to-point speech software that
accumulates a histogram based on the difference between time stamps on
arriving packets and a local clock that is initialized to the time stamp of
the first packet to arrive. Time resolution (histogram cell size) was
22.5 ms, the parcel time in the PCM speech encoder in the PVT. A traffic
background was provided by having other terminals engaged in similar loop
experiments at the same time that a measurement was being made. For an
experiment with two 64-kbps speech streams of traffic, the results shown
increasing delay and dispersion of delay as the loop point moves farther away
from the PVT. When the packets go only to the LEXNET, they all arrive in the
first 22.5-ms time slot. When they go to the gateway and back, almost all
return within two intervals (45 ms). Going into the Packet Satellite
Interface Message Proc^ssor (PSAT) and back adds another 22.5 ms unless
packet aggregation ⸭ used. With our normal aggregation interval of 44 ms,
the delay in coming back from the PSAT increases to 180 ms, and reaches a
value of 540 ms when the packets are sent to the satellite. These numbers

14

are for the packets that experienced the greatest delay in each case. For
the satellite loop the dispersion extended across six cells (roughly 140 ms
of dispersion). The results are in general agreement with our expectations
for the conditions that were tested.

117953-N

TO/FROM SATELLITE

Ⓕ

LOOP POINTS

Ⓔ

ESI

A. LEXNET

B. GATEWAY

C. NULL PSAT
   (DH Interface)

PSAT   Ⓓ

D. PSAT
   (Host Module)

Ⓒ

E. IF OUTPUT
   OF EARTH STATION
   INTERFACE (ESI)

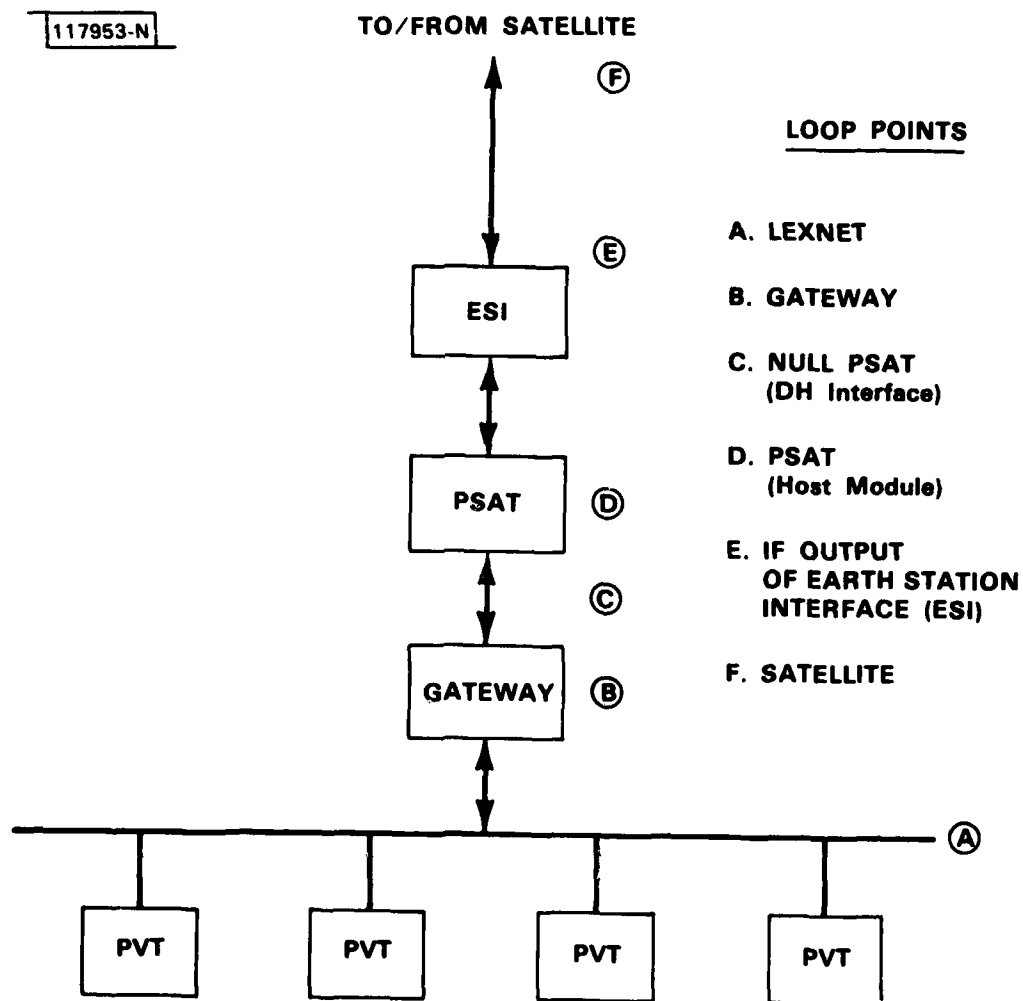GATEWAY   Ⓑ

F. SATELLITE

Ⓐ

PVT   PVT   PVT   PVT

Fig. 3. Loop delay measurement configurations.

III. MINICONCENTRATOR

A. HARDWARE STATUS

In the last report we noted a problem with the UMC-Z80 hardware which prevented using timer interrupts from the UMC processor to control the packet dispatch rate. The trouble occurred when the UMC-Z80 interrupt coincided with an interrupt from the internal 60-Hz clock in the PDP-11/44. This problem has now been fixed by the most recent engineering changes carried out by Associated Computer Consultants, the UMC-Z80 manufacturer. We are now able to run the gateway with timer interrupts from the UMC at the intended rate, which allows us to dispatch stream messages to the PSAT every 44 ms. All the UMC boards in the miniconcentrator at all WB SATNET sites have now been updated to include this change. The new stream packet dispatch interval is now well matched to the PSAT stream interval of 42.4 ms. The previous dispatch interval of 50 ms, which was controlled by the 60-Hz clock in the PDP-11, wasted about one-third of the stream capacity.

B. GATEWAY SOFTWARE

The miniconcentrator is fully operational as an internal gateway supporting both Internet Protocol (IP) and stream (ST) functions, and has been used as a central element in packet speech experiments on the WB SATNET. The gateway includes all basic IP and ST functions for point-to-point calls, and now includes an ST conference-handling capability. PDP-11-based gateways are operating at ISI, DCEC, SRI, and Lincoln. Two PDP-11 gateways are operating at Lincoln. Gateway software is operating to support connections to the WB SATNET, LEXNETs, the ARPANET, the Exploratory Data Network (EDN) at DCEC, and circuit-switched Telephone Office Emulator. We expect that software to support Packet Radio net connections will become operational during the next quarter. The gateway software is modularized for easy generation of N-headed versions (i.e., simultaneous connection to N nets). It employs a multiprocessor design (PDP-11 central processor, with UMC-Z80 servicing each connected net); and the software is implemented as a

17

multiprocess structure. Control and monitoring is provided via a TTY connection to a support processor, and support software is operational under UNIX versions 6 and 7 as well as under EPOS. A background timer process is available for performance measurements.

During this reporting period, the miniconcentrator gateway has been a key element in several milestone experiments on the WB SATNET. These include: (1) the 18 November LEXNET-to-LEXNET speech tests with ISI; (2) a subsequent test including use of the ISI-developed switched-telephone network interface (STNI) board to connect from a PVT to an ordinary phone; (3) three-site tests on 4 March involving simultaneous stream communication among Lincoln, ISI, and DCEC; and (4) demonstration on 2 February of speech compatibility between the miniconcentrator and the BBN Voice Funnel (using voice packets transmitted as ST packets encapsulated in IP, since ST is not yet implemented in the Funnel).

Specific developments in the gateway software accomplished during this reporting period are summarized below.

The capability to aggregate a number of ST packets into an aggregated packet for the transmission over the WB SATNET has been completed and checked out. This capability has been crucial in terms of decreasing the load on the PSAT in terms of the number of packets per second transmitted. Packet aggregation was utilized in all the milestone speech experiments described above.

An ST source-routing mechanism was implemented in which the gateway could be requested to send the ST data packets on a loop through any network before sending the packets to their final destination. This provided a debugging environment in which local LEXNET PVT to PVT connections could be looped through the PSAT. A number of important tests and network measurements have been carried out using this ST source-routing mechanism.

The implementation of an ST conference-handling capability was initiated during the past quarter. When an ST conference connect call comes in, the gateway queries the conference access controller (AC) running on a LEXNET PVT about the conference. From the information provided and subsequent

conference control messages, the gateway builds and modifies its internal
tables in order to handle the conference. Later, when conference speech
packets arrive, the gateway manipulates the forwarding bit maps in order to
forward the packets to the appropriate destinations. For nets with a
broadcast capability, the gateway achieves forwarding through use of
broadcast addressing rather than packet replication. To date, a conference
between two participants on different LEXNETs has been achieved through the
gateway. Further development and test of conferencing is expected to be a
major focus of effort during the next few months.

Commands have been provided to query the gateway status and to modify
its execution modes. Among recent commands have been those for creating and
deleting PSAT streams, selecting the network to be used for ST source
routing, shutting down or pausing the UMC-Z80 programs, and printing network
specific message counts or miscellaneous status information. The
organization of the gateway program permits relatively easy incorporation of
additional commands as the need may arise without adversely affecting the
foreground functions of sorting, routing, and dispatching the speech traffic.

An ST gateway-to-gateway protocol is being implemented in order to allow
utilization of the broadcast capability of the PSAT network. Currently, the
gateways have a mechanism for exchanging ST HELLO messages with each other so
that each gateway can thereby know which other gateways are in operation.
This exchange of information is being extended to provide for the creation of
PSAT groups and the subsequent joining of the groups by the gateways
involved. The group addresses will be used as a means for broadcasting
conference messages.

External gateway control has been provided for controlling the size-
related parameters involved in the request for PSAT stream reservations.
This has enabled us to bring up gateways at three sites (Lincoln, ISI, and
DCEC) and establish simultaneous communication between combinations of these
sites using PSAT streams. In this experiment LEXNETs were employed at
Lincoln and ISI and a Lincoln-developed Telephone Office Emulator (TOE)
connected through a Packet-to-Circuit-Interface (PCI) was used at DCEC. (The
TOE and PCI were developed under DCA sponsorship.)

19

The new Host Access Protocol (HAP) addressing was incorporated in the gateway to coincide with corresponding addressing changes in the PSATs, which were introduced in March.

The "downloader" program was further modularized so that additional versions could be generated. Currently, there are versions which run as user programs in Version 7 (V7) UNIX at Lincoln, V6 UNIX at DCEC, and EPOS at ISI. Upcoming is a version that will run at SRI under their V7 UNIX. The main control module of this program is common to all versions; unique modules for performing I/O, system calls, and handling the vagaries of the various operating systems provide for the different versions.

A command was added to the downloader to accept a file containing a dialogue between a user and the downline computer; the dialogue is played out until it is complete or until the response is in the dialogue file. This mechanism provides for quick error-free booting of gateways and their attendant UMC-Z80 control programs.

C. MULTIPORT BUFFER MEMORY

The major portions of the Multiport Buffer Memory (MBM) design have been completed and a Technical Report describing the design is in preparation. A brief summary of the design is given here.

The MBM is a shared memory resource which is attached to a complex of microprocessor-based subsystems to provide a flexible buffer for packet traffic in the miniconcentrator environment. The design provides for the shared use of a 64K x 16 Main Memory array among up to four attached UMC-Z80 processor ports. The Multiport Buffer Memory is designed to improve the throughput and efficiency of packet traffic through the miniconcentrator. It accomplishes this by providing a common pool of buffer memory management which incorporates an intelligent controller to relieve each UMC-Z80 and the miniconcentrator of the overhead associated with packet storage and forwarding. The basic scenario for packet communication using the MBM involves the assignment of parameters which define the number of words and the starting location of data packets stored in the shared memory. These

parameters are attached as parts of the header to packets which are circulated in the network. A memory management strategy is employed to keep track of the manner in which packets are stored in the memory such that the details of the stringing of 16-word data blocks is transparent to each port attached to the MBM.

The Main Memory elements are 64K x 1 dynamic RAMS (MCM6664). Tests of these components conducted over a period of several weeks produced no errors and this experience coupled with recently published reports of low soft error rates for these devices led to a decision to avoid the added complexity of employing error detection and correction in the Main Memory array.

The block diagram for the Multiport Buffer Memory is shown in Fig. 4. The basic architecture consists of an external UMC-Z80 bus attached to a set of Input and Output Latches which connect to the main data bus of the MBM. Each external bus is connected to individual complexes which consist of Z80DMA, Z80PIO, and address decode logic. Data exchanges between individual UMC-Z80 processors and the MBM are conducted via DMA transfers or under control of input and output instructions in the UMC-Z80s. Control and status information is passed through the PIOs. The Decode Logic is used to decode reserved addresses on the UMC-Z80 address bus in order to steer data on the UMC-Z80 data bus to either the DMA, the PIO, or the Input or Output Latches.

The Input and Output Latches consist of latch pairs which are alternately selected to accomplish the conversion from the 8-bit bytes on the UMC-Z80 busses to the 16-bit word structure of the Multiport Buffer Memory. The 16-bit word size was selected to allow sufficient time between successive accesses to allow fully interleaved operation of the four ports in a manner transparent to full DMA operation on each individual port. This results in an access time for the entire complex of 625 ns for either a read or write operation which is well within the limits of the MCM6664 access specification of 350 ns.

The basic architecture of the memory complex consists of a tri-stated two-bus structure to which a set of registers is attached for memory management and bus arbitration functions. The basic control element for the

21

architecture resides in a Finite State Machine which provides control signals to transfer data between the various registers attached to the data and address busses. The control of these transfers is effected by the Finite State Machine which is implemented as a set of PROMS which are addressed by a combination of new input parameters and previous state values. This design was selected in place of a microprocessor controller because investigation of available microprocessors (either bit slice or single chip) revealed that none of the devices available would allow full control of the architecture within the 625 ns allotted to each port of the overall system. The Finite State Machine is, in effect, a special-purpose microprocessor which is programmable by virtue of changes in the contents of the PROMS.

The memory management philosophy consists of conceptually partitioning the Main Memory into blocks of 16 words which are addressed by an Auxiliary Pointer Memory which maintains a string-structured list of 16-word blocks in Main Memory. The pointer memory is composed of an array of 4 static RAM elements (IMS1420- 4K x 4 RAMs) which provide 4K x 16 bits of pointer memory. The address pointer data comprises 12 bits which point to the 4K blocks of 16 words in Main Memory. The remaining 4 bits are used for parity, end-of-string, and buffer-full flags, and a spare. Parity is used in connection with this memory because of its critical role in the memory management scenario. A detected parity error is reported to the granted UMC-Z80 via an interrupt through its PIO and the operation is aborted before it can affect the pointer chain.

An address Register File/Counter consisting of a 4 x 16 Register File and an Address Counter provides addresses for both the Main and Auxiliary Memories. The four-deep Register File maintains the status of the address between access grants to different ports. A multiplexed input to the Address Register File allows the counter to increment the address between successive accesses without having to compete for the data bus within an ongoing memory cycle. A similar Length Register File/Counter complex is used to determine the completion of read or write operations by detecting the end of the stored or retrieved packets. These end-of-packet detections are reported to the

22

appropriate PIOs for notification of the UMC-Z80 processors. A special "Free List Register" is situated between the data and address busses of the main architecture in order to provide a running pointer to free blocks in Main Memory.
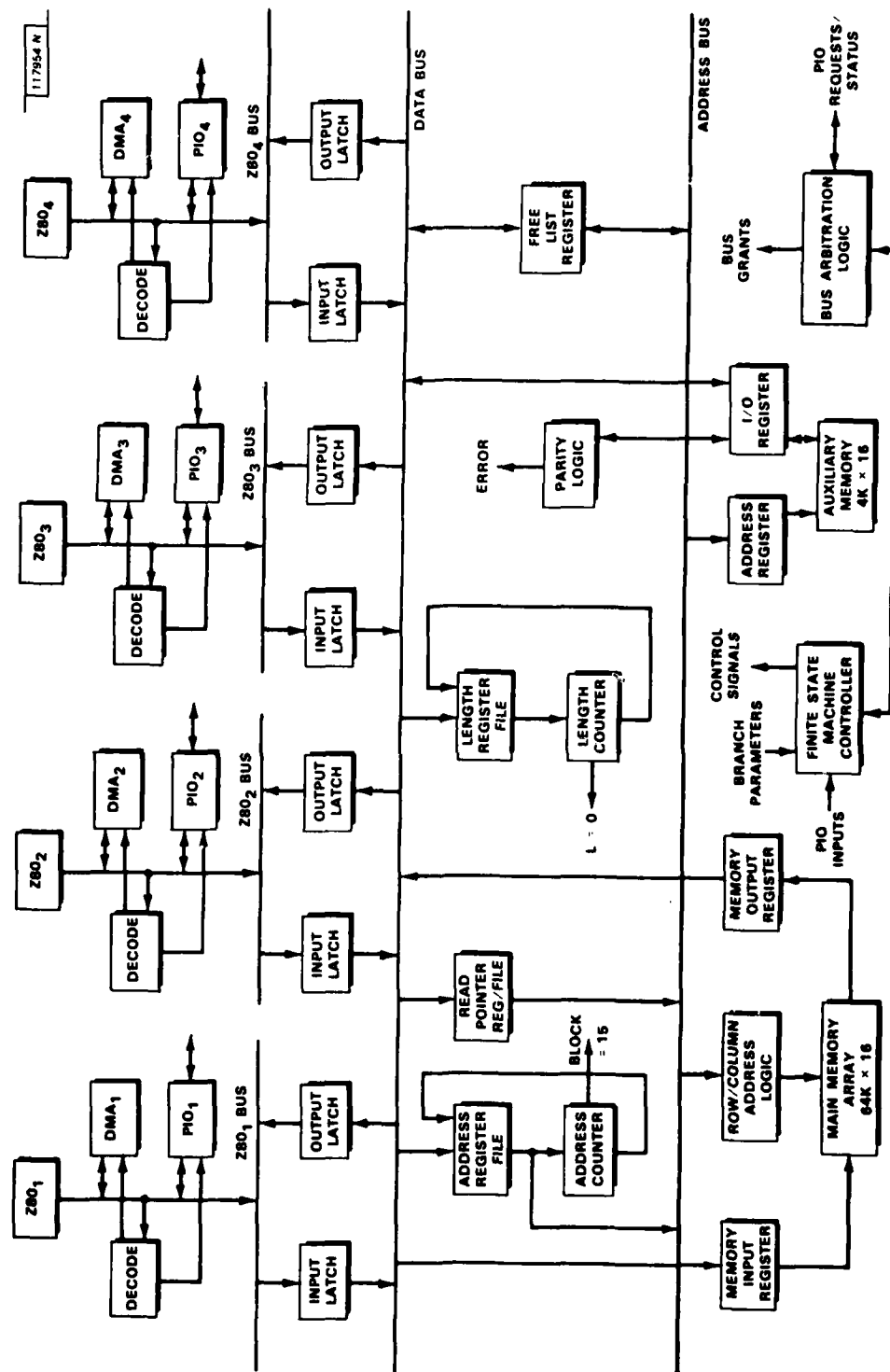
Fig. 4. Data/address bus architecture.

117954 N

IV. EXPERIMENT PLANNING AND COORDINATION

Lincoln has continued to perform a lead role in the coordination of efforts aimed at achieving a stable operational status for the WB SATNET. We have coordinated and participated in numerous milestone experiments on the WB SATNET, but steady multisite system operation has not yet been achieved. At the 1-2 March Wideband Meeting, Lincoln presented an overview of system problems in the WB SATNET, and a set of recommendations to correct the problems. In summary, it was reported that no single subsystem could be targeted as the primary source of difficulty. Rather, small subsystem bugs in each component, as well as numerous equipment failures and loss of time during repairs, have combined to prevent us from achieving steady multisite operations for extended periods of time. The Lincoln recommendations, many of which have been implemented or are in process, include: (a) fix all known subsystem problems; (b) develop and exploit the involvement of the Network Operations Center at BBN; (c) develop an effective arrangement for spare equipment and quick equipment repair; and (d) schedule and execute numerous tests and demonstrations to exercise the system.

The most pressing need at this time is to exercise the system in a near-operational sense as much as possible. To this end a schedule has been devised and agreed upon for certain time periods two days per week when the satellite channel will be available for general coast-to-coast speech experimentation, leaving the rest of the week for dedicated test and development activity coordinated by BBN. As a result of this extended near-operational activity it is expected that any remaining system problems will be identified and resolved.

Plans are in progress for instrumentation which will allow continuous operation and monitoring of multiple simultaneous speech channels, both loopback and cross-country, to compile detailed information about the system performance on an extended basis including nights and weekends. At present it appears that these functions can be performed by suitably modifying LEXNET traffic emulator software to detect and report packet loss percentages and interruptions of connections. The philosophy in this mode is to regard the

25

WB SATNET as a transmission medium to be evaluated from a speech user point of view, without directly involving PSAT self-diagnostics such as packet error counts and other trap messages. A more complete characterization of the channel will result from independently and simultaneously collecting information from both kinds of sources.

The primary current focus of our coordination activities is preparation for a packet speech demonstration in June involving four sites connected through the WB SATNET. The equipment configuration for this demonstration is shown in Fig. 5. Plans call for demonstration of point-to-point and conferenced speech, using PCM, ECVSD, and LPC coding. A key aspect of the configuration is an Internet connection between LEXNETs at Lincoln and ISI, over the WB SATNET, to a PRNET at SRI. Much of the basic speech capability shown in Fig. 5 is currently in place. The primary goal which must still be achieved to support the demonstration is reliable and steady multisite operation on the WB SATNET.

An analysis of Wideband Network throughput capabilities and limitations has been carried out, and was presented at the Wideband Meeting and published as a Wideband Working Note. A model for the network was established, in which six choke points were identified in such a way that their effects on system throughput could be individually evaluated. Two scenarios were chosen as benchmarks for the analysis. In the first scenario there are two sites with identical LEXNETs, and the object is to find the largest possible number of point-to-point calls that can be supported with identical voice terminals at the two sites. In the second scenario there are N identical sites, each supporting N calls (one to each other site, and one conference call including all the sites), and the object is to maximize the value of N. Equations were derived for each of the choke points in the network, permitting calculation of the desired maxima as functions of various parameters such as interface bit rates, voice coder bit rates, packets-per-second processing capability, and so on. Tables of values were computed for each scenario, for three different voice code rates (2.4, 32, and 64 kbps). A brief summary of some of the key values from these tables was presented at the Wideband Meeting.

The Wideband Meeting, held at DARPA on 1-2 March, was attended by approximately 24 people representing all the organizations participating in the project. The meeting opened with comments by each of the sponsors (DARPA and the DCA) about recent progress, perceived problems, and both near- and long-term goals. Each organization gave a presentation on topics of interest, and there was a large amount of discussion. A topic of special interest was the performance of the satellite network itself, both in terms of the throughput issues mentioned above and in terms of the reliability and general performance problems of the earth stations. One outcome of this part of the discussion was a meeting held some three weeks later at Western Union, attended by representatives of DARPA, DCA, and Lincoln, at which plans were established for correcting each of the Western Union problems that had been identified. A general summary of the Wideband Meeting is in preparation, and will be distributed to each of the attending organizations.
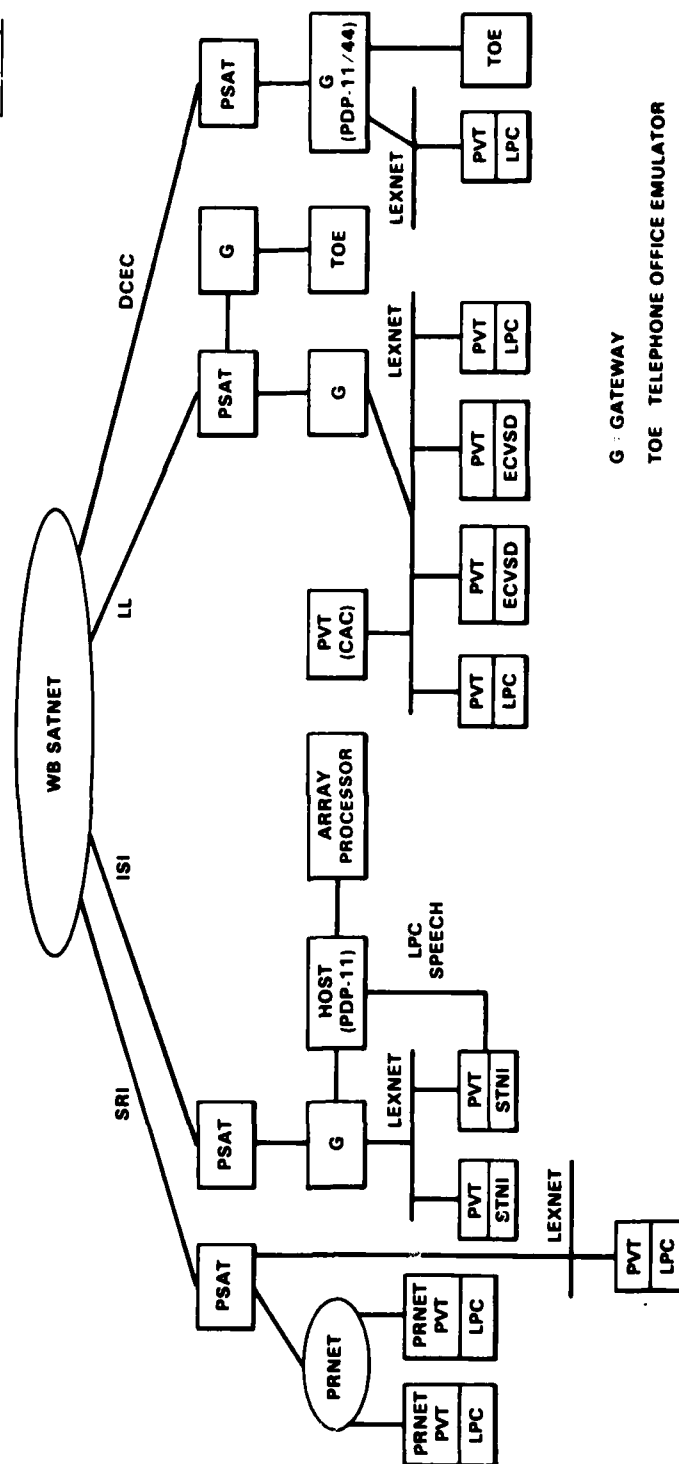
Fig. 5. Four-site packet speech demonstration configuration.

G - GATEWAY

TOE TELEPHONE OFFICE EMULATOR

CAC CONFERENCING ACCESS CONTROLLER

117955-N

V. VOICE CONTROL OF NETWORK VOICE CONFERENCING

A system design and a set of experiments has been defined to lead to a capability for setting up conference calls in the wideband network environment via voice commands. The experiments require a voice control operator (VCOP) which includes (1) a word recognition system to identify spoken words; (2) a speech synthesizer to produce speech feedback, prompts, and acknowledgments; (3) interfaces from the recognizer and the synthesizer to a packet voice terminal; and (4) software to control the recognition system and the synthesizer. A block diagram of a voice control operator is presented in Fig. 6. In this diagram the packet voice terminal (PVT) provides audio I/O and an NVP interface to the Wideband Internet via a LEXNET. It also controls the recognition system and synthesizer. The recognition system provides real-time recognition of words in the audio output from the PVT. The synthesizer produces audio prompts and acknowledgments which are fed into the audio input of the PVT.

The Threshold Technology T580 speaker dependent word recognition system and a VOTRAX VSM/II Speech Synthesizer were selected for use in the VCOP. These are commercial devices which can be controlled via characters sent over an RS232 link. The T580 recognition system was chosen because it provides very high recognition accuracy when used to identify isolated words produced by individual talkers but is available at much lower price than either of the two other acceptable systems. These systems, the Verbex 1800 and the Nippon Electric DP-100, cost more than five times as much as the T580 ($65,000 vs $12,000). The VOTRAX VSM/II synthesizer was chosen because of its flexibility and low cost. It includes a prestored vocabulary of 1300 words, can operate in a phoneme-based mode which allows an unlimited vocabulary, and costs roughly $1,000. The T580 system has been received and tested under local control from a CRT terminal. The electrical interface between VCOP components has been constructed and tested. It allows the synthesizer and recognition system to be controlled either from a PDP-11/44 or from a PVT terminal.

29

The major components of the VCOP control software have been identified and the structure of this software has been specified. Programs are being written in C on a PDP-11/44 under a UNIX operating system. In the initial stages of VCOP development this 11/44 will be used to control the synthesizer and recognition system, to permanently store reference templates needed by the word recognition system, to structure the conversation between the caller and the VCOP, and to instruct the PVT to send NVP tokens.

Initial program development has focused on controlling the synthesizer and on developing the logic to control a conversation between a caller and the VCOP. We have written a phoneme editor which allows phrases to be constructed using words and phonemes instead of using the hex codes required by the VOTRAX synthesizer. Phrases have been created using this editor, downloaded to the VOTRAX synthesizer using the UNIX system commands and a C program we have written, produced by the VOTRAX synthesizer, and then modified after listening to increase naturalness and intelligibility.

Our experience has been that although many persons can immediately understand the output of the VOTRAX synthesizer, many cannot. These persons require training to identify words and phrases. The VOTRAX synthesizer will be adequate in voice control experiments as long as callers have had previous exposure to VOTRAX-produced speech. In a more advanced VCOP it would be preferable to store LPC parameters describing phrases produced by a talker and to download these parameters to a synthesizer to produce phrases.

A scenario which describes the interactions between the conference call initiator, conference participants, and the VCOP has been written. An example of the dialogue produced during these interactions is presented in Fig. 7. We are currently writing a program to control the VCOP which supports these interactions. Prompting and response phrases will be produced using the VOTRAX synthesizer and input from a caller will initially be entered on a CRT terminal keyboard. The keyboard will be eliminated and replaced by the T580 re ognition system after control programs have been written for the T580.

30

Preliminary modifications to the NVP protocol which allow a conference call to be set up and controlled using a voice control operator have been defined. These modifications describe how the voice control operator provides the conference controller with information about the conference. They also describe how the conference controller uses the voice control operator to bring new participants into a conference and to notify the moderator when a participant leaves or joins the conference.
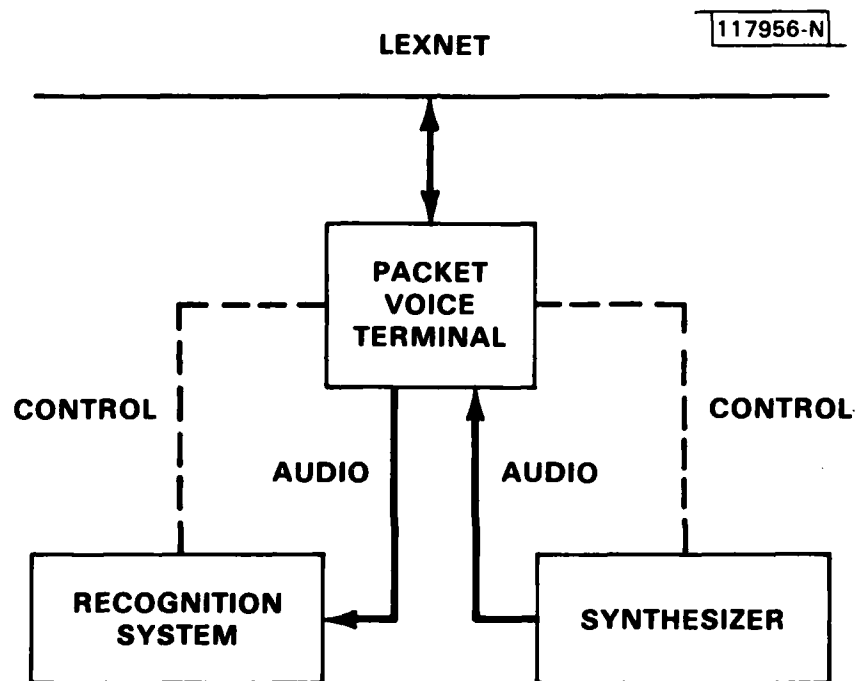
LEXNET 117956-N

Fig. 6. Block diagram of voice control operator (VCOP).

31

INITIATOR DIALS VCOP

VCOP	"HELLO MR. FORGIE,"
	"WOULD YOU LIKE TO SET UP A CONFERENCE?"
INITIATOR	"YES."
VCOP	"WHO SHOULD BE IN THE CONFERENCE?"
INITIATOR	"GERRY"
VCOP	"GERRY O'LEARY"
INITIATOR	"CLIFF"
VCOP	"CLIFF WEINSTEIN"
INITIATOR	"DONE"
VCOP	"WHAT IS THE CONFERENCE TOPIC?"
INITIATOR	"VOCODERS"
VCOP	"VOCODERS?"
INITIATOR	"YES"
VCOP	"PLEASE HOLD"
	"I AM SETTING UP THE CONFERENCE."

VCOP DIALS GERRY O'LEARY'S PVT

VCOP	"IS THIS GERRY O'LEARY?"
G.O.	"YES"
VCOP	"WOULD YOU LIKE TO JOIN A CONFERENCE ON VOCODERS WITH JIM FORGIE AND CLIFF WEINSTEIN?"
G.O.	"YES"
VCOP TO INITIATOR	"GERRY O'LEARY JUST JOINED."

ETC.

Fig. 7. An example of the dialogue between a conference call initiator and the VCOP and b▪ ▪een a conference participant (G.O.) and the VCOP.

32

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
ESD-TR-82-052	AD-A120 433	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
Packet Speech Systems Technology	Semiannual Technical Summary 1 October 1981 — 31 March 1982
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
Clifford J. Weinstein and Peter E. Blankenship	F19628-80-C-0002

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Lincoln Laboratory, M.I.T. P.O. Box 73 Lexington, MA 02173-0073	ARPA Order 3673 Program Element Nos. 61101E/62708E Project Nos. 2D10/2T10

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
Defense Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, VA 22209	31 March 1982
	13. NUMBER OF PAGES
	42

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report)
Electronic Systems Division Hanscom AFB, MA 01731	Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

packet speech	ARPANET	LEXNET
network speech	SATNET	digital channel vocoder
voice conferencing		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report describes work performed on the Packet Speech Systems Technology Program sponsored by the Information Processing Techniques Office of the Defense Advanced Research Projects Agency during the period 1 October 1981 through 31 March 1982.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 Jan 73